

# Тонкая настройка трансляции у бактерий

## Comparing sequence evolution of homomers versus heteromers

Julliane Negre, Laura Avinyó

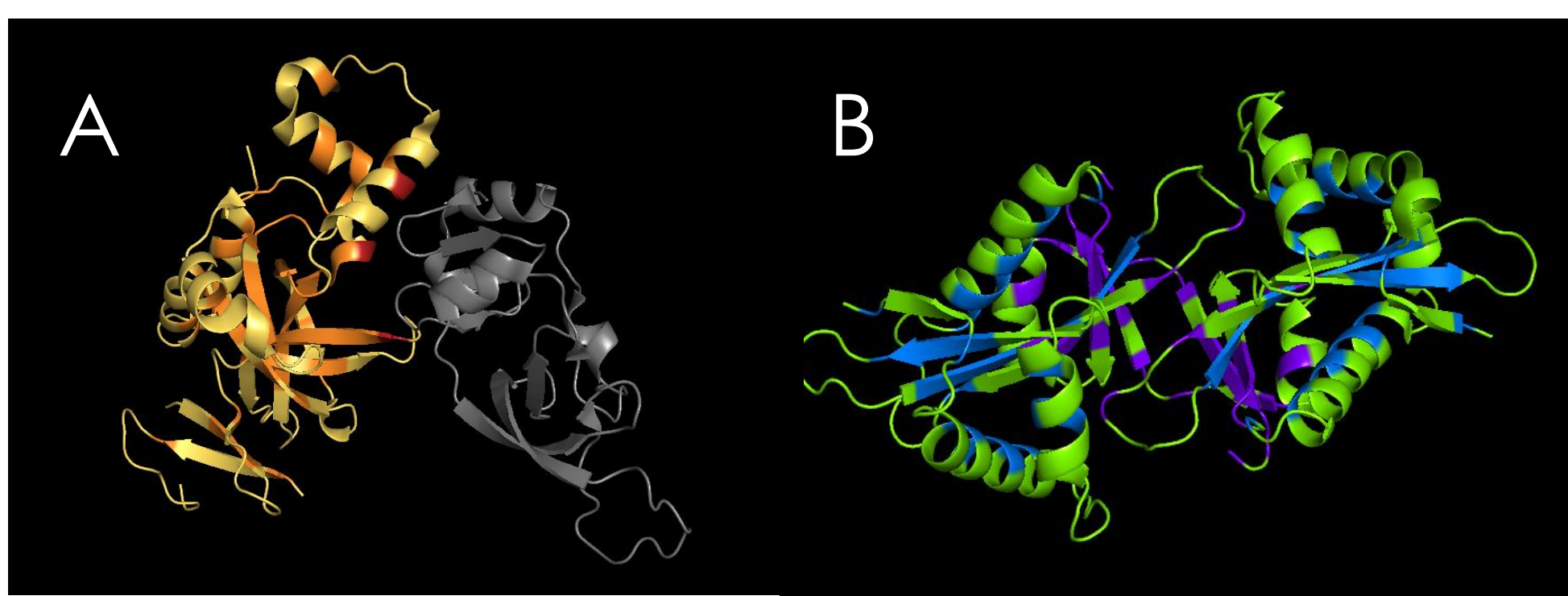
### Background

It is well known that proteins form complexes to increase their stability and improve functionality. There are mainly two types of complexes — homomers, in which all subunits are identical, and heteromers, in which subunits are encoded by different genes. However it is still not well understood whether these types have differences in the mode of evolution. Here we address this question studying five homomer and five heteromers.

Хорошо известно, что белки формируют комплексы, что улучшает их стабильность и функциональность. Существуют два основных типа комплексов: гомомеры, в которых субъединицы идентичны, и гетеромеры, в которых субъединицы кодируются различными генами. Однако неясно, есть ли различия в эволюции гомомерных и гетеромерных субъединиц. В настоящем исследовании сопоставлена эволюция пяти семейств гомомеров и пяти семейств гетеромеров.

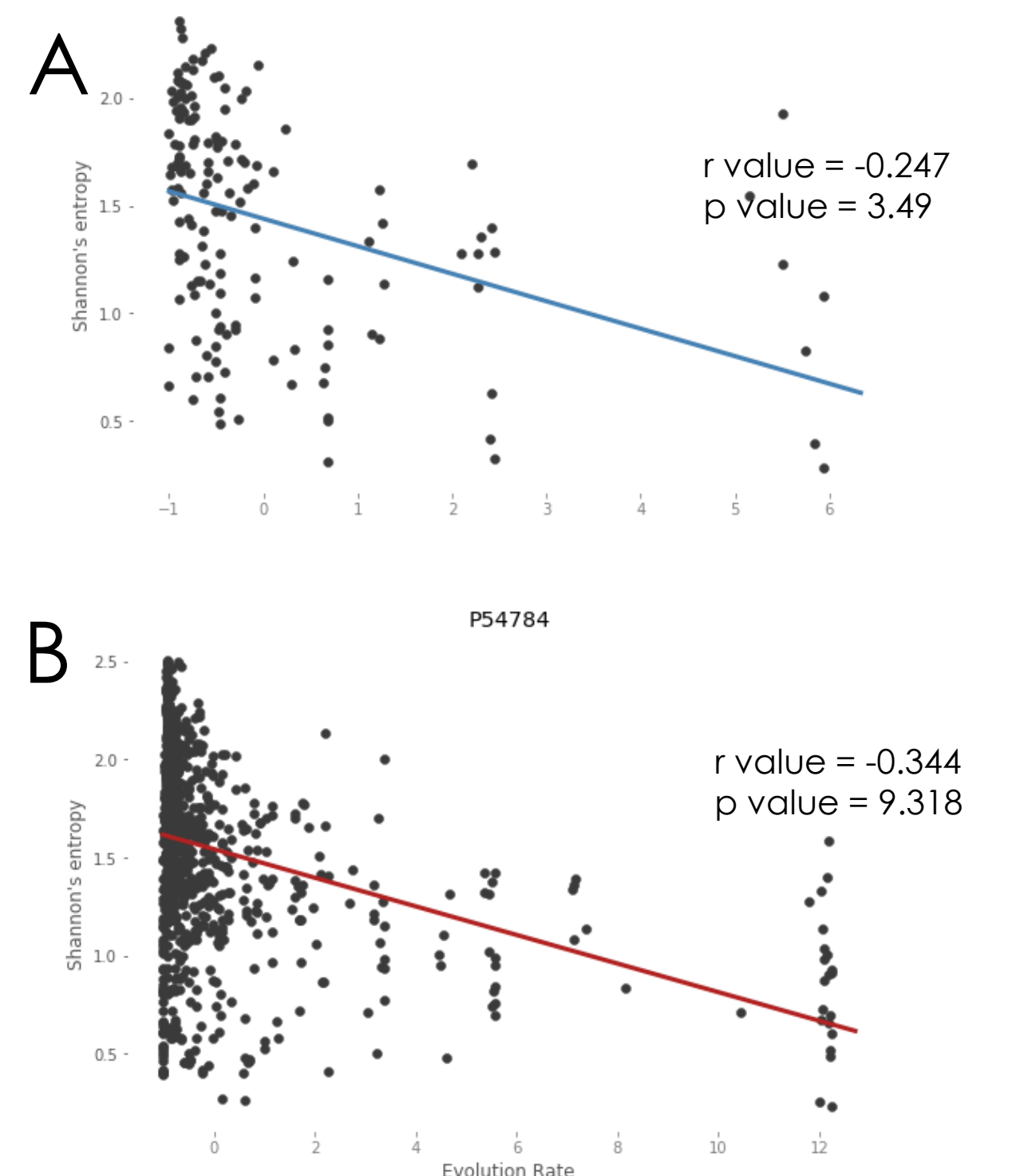
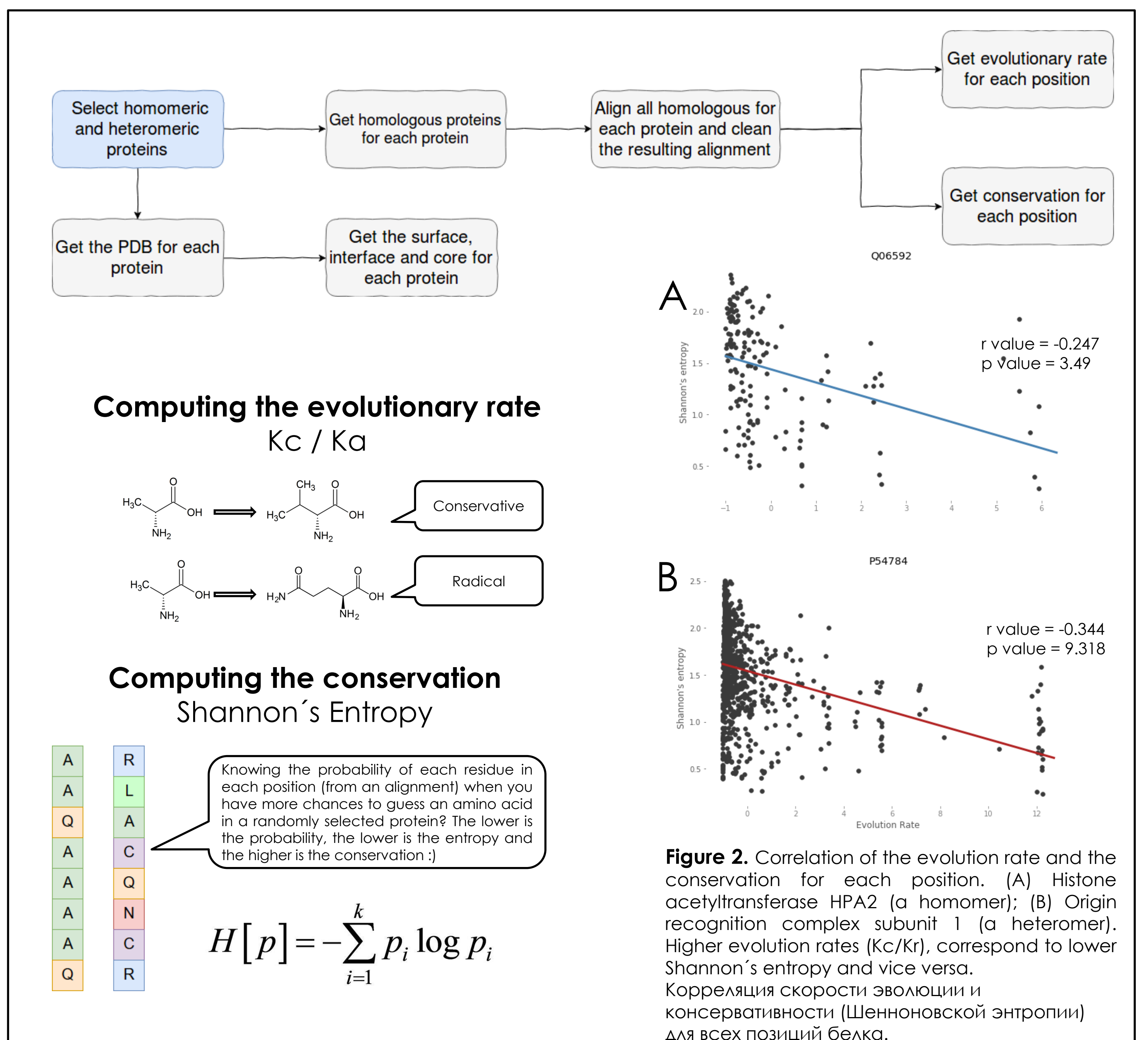
Ang mga protina ay bumubuo ng iba't ibang complexes upang tumaas ang kanilang estabilidad at kakayahan. Mayroong dalawang mas kilalang complex — homomers, na binubuo ng mga kasangkap na magkapareho, at heteromers na mayroon kasangkap na gawa ng iba't ibang genes. Ngunit, hindi pa masyadong naiiindihan kung mayroong pagkakaiba ang ebolusyon ng mga complex na ito. Dito natin maisasagot ang tanong na ito sa paraang nang pag-aaral ng limang homomer at limang heteromer.

Es coneix que les proteïnes formen complexos per millorar la seva estabilitat i funcionalitat. Hi ha dos tipus de complexos principals: els homòmers, en els quals les subunitats són idèntiques, i els heteròmers, en els quals les subunitats estan codificades en gens diferents. Tot i això, encara no és entès del tot si aquests tipus tenen diferents maneres d'evolucionar. Aquí adrecem aquesta qüestió estudiant cinc homòmers i cinc heteròmers.



**Figure 1.** (A) 3D representation of P54784, a heteromer subunit. Red — the interface with other subunits, yellow — the surface, orange — the core. (B) 3D representation of Q06592, a homodimer. Purple — the interface, green — the surface, blue — the core. Структура гетеромера (A) и гомомера (B). Показаны поверхность контакта, другая поверхность и кор.

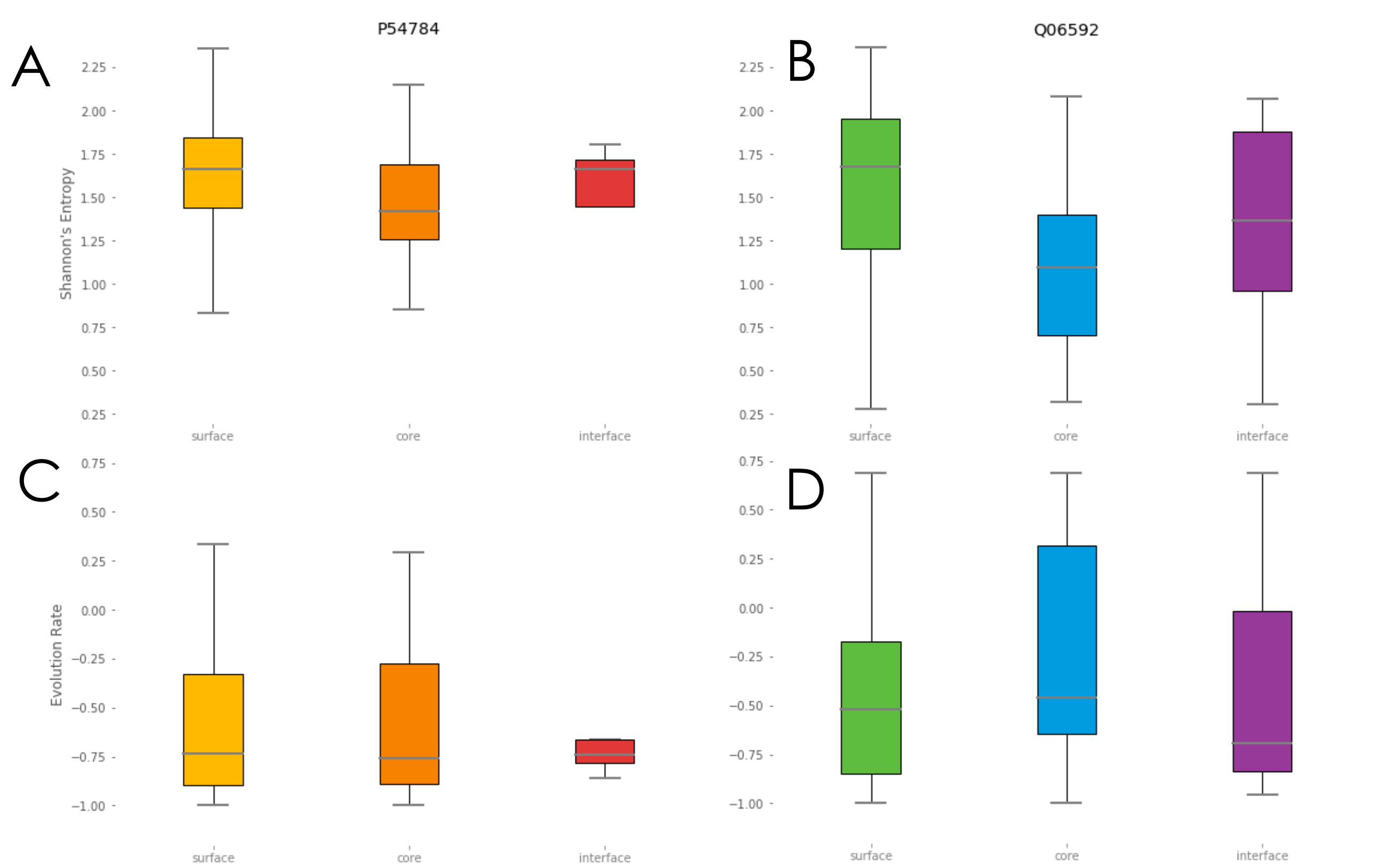
### Workflow and Background



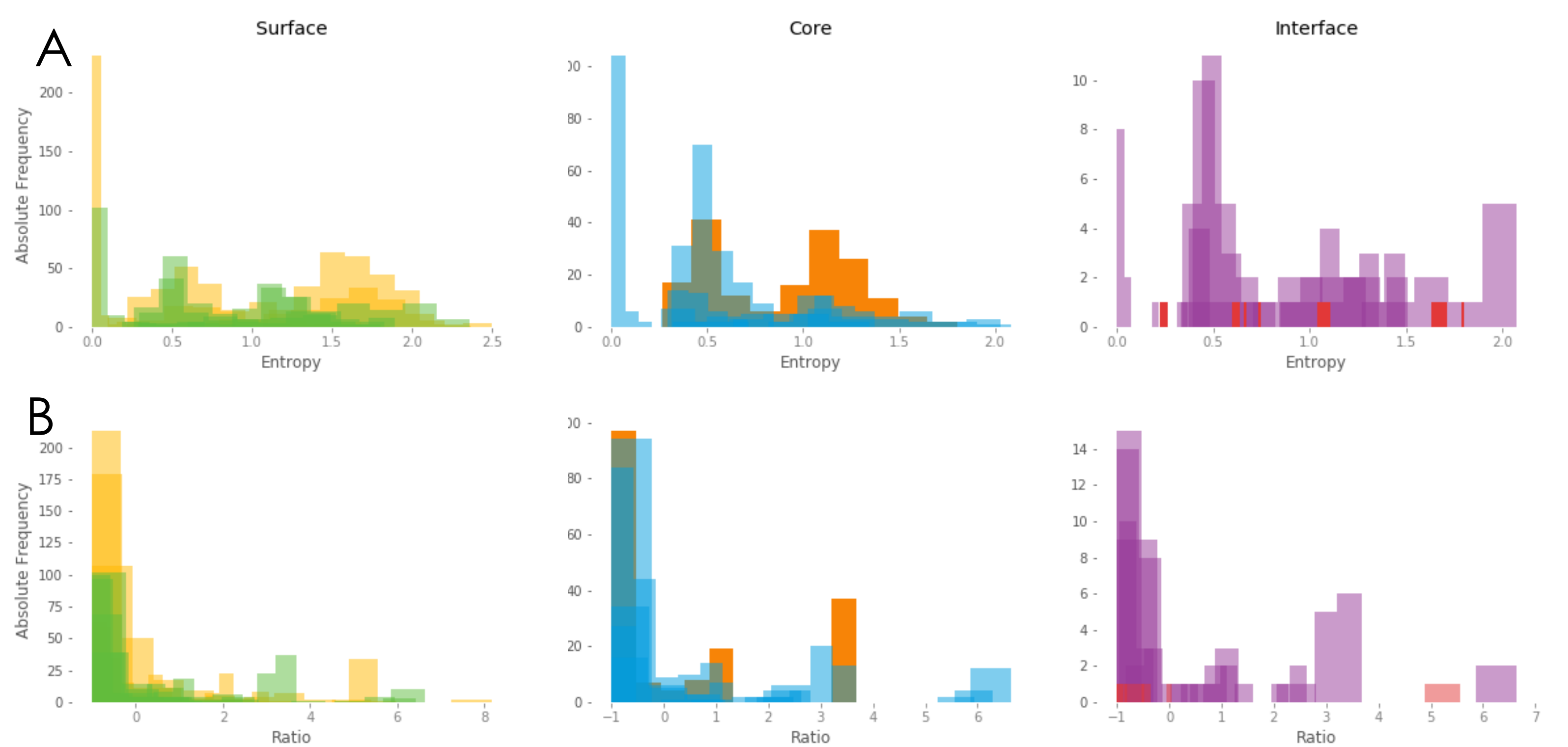
**Figure 2.** Correlation of the evolution rate and the conservation for each position. (A) Histone acetyltransferase HPA2 (a homomer); (B) Origin recognition complex subunit 1 (a heteromer). Higher evolution rates ( $K_c/K_r$ ), correspond to lower Shannon's entropy and vice versa. Корреляция скорости эволюции и консервативности (Шенноновской энтропии) для всех позиций белка.

### Results

#### Measures of Shannon's Entropy and Evolutionary Rate per structural property



**Figure 3.** Distribution of the evolutionary rate and conservation of positions in each structural category. A and C show Origin recognition complex subunit 1. B and D show Histone acetyltransferase HPA2. Note that the core is more restricted (low Shannon's entropy, high  $K_c/K_r$ ) than the interface in both proteins. In our data it happens for 80% of homomers and 60% of heteromers. Распределение скорости эволюции и консервативности для каждой структурной категории.



**Figure 7.** Histograms per structural category with comparisons between heteromers and homomers. (A) The distributions of entropy in the surface and core for heteromers (resp., yellow and orange) and homomers (resp., green and blue) are similar. (B) The distributions of evolution rates do not show any significant difference. Сравнение скоростей эволюции и Шенноновской энтропии в гомомерах и гетеромерах не выявляет значимых различий.

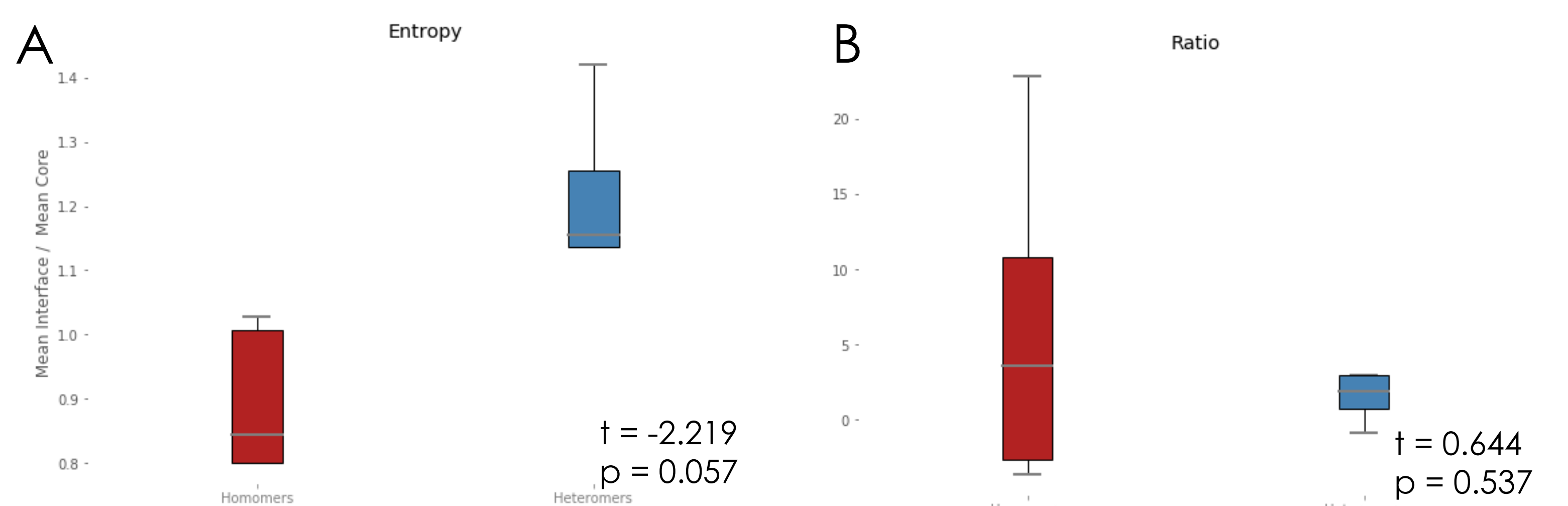
### Conclusions and future work

In summary, a comprehensive pipeline was constructed. The pipeline calculates Shannon's entropy and Evolution Rate, and normalizes these values for any given set of protein homologs. The pipeline also assigns structural categories (surface, core, and interface) to amino acids of the query proteins.

Conservation (Shannon's entropy) and Evolutionary rates ( $K_c/K_r$ ) are correlated. The core is the most conserved region. Homomer subunits tend to have a larger surface area that interacts with the other subunits as compared with heteromers. Homomers tend to have a higher restriction in the evolution of their interfaces, however we do not have sufficient data to prove statistical significance of this observation. Finally, the evolutionary rate does not differ between homomers and heteromers.

The study may be extended using the created pipeline for a larger data set. Additional analysis may be done using other categorizations of amino acids as a parameter for a radical or conservative change.

All data and the code are available at [https://github.com/LauAvinyo/SMTB18\\_Project](https://github.com/LauAvinyo/SMTB18_Project).



**Figure 9.** Distributions of normalized entropy (A) and normalized evolutionary rate (B). The entropy plot shows that homomers are more conserved than heteromers, this difference is not statistically significant (probably due to the small sample size). The evolutionary rate plot shows no difference between homomers and heteromers. Распределение энтропии показывает большую консервативность гомомеров по сравнению с гетеромерами однако различие статистически незначимо (возможно из-за небольшого размера выборки). Распределение скоростей эволюции не выявляет различий.